JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 40, 1-11 (1972)

Envelope Programming and a Minimax Theorem

D. J. WHITE

Department of Decision Theory, Faculty of Economic and Social Studies, University of Manchester, Manchester, England

Submitted by G. Leitmann

1. STATEMENT OF PROBLEM

Consider the following problem:

$$\max_{x \in R}[\phi(x)], \tag{1}$$

where $\phi(x)$ is a strictly concave function of x, and R is convex and a subset of n dimensional Euclidean space E_n.

Let x be any point in R. Then $\phi(x)$ can be represented by the minimum value of the hyperplanes to ϕ taken over all $t \in E_n$. Writing this more formally we have

$$\phi(x) = \min_{t \in E_n}[\phi(t) + (x - t)' \nabla\phi(t)], \tag{2}$$

where

$$\nabla'\phi(t) = \left(\frac{\delta\phi(t)}{\delta t_1}, \frac{\delta\phi(t)}{\delta t_2}, \dots, \frac{\delta\phi(t)}{\delta t_n}\right). \tag{3}$$

This is easily proved by differentiating the right side of Eq. (2) and using the fact that $\nabla^2\phi(t)$ is negative definite for all $t \in E_n$ (in fact we need only confine our requirements to a sufficiently large but bounded subset of E_n), where

$$\nabla^2\phi(t) = \left(\frac{\delta^2\phi(t)}{\delta t_i\, \delta t_j}\right). \tag{4}$$

We are thus expressing the function $\phi(x)$ in terms of its tangent hyperplanes and it is seen to be the "envelope" of these hyperplanes. We thus devise the term "envelope programming."

Let us now define the function $\theta(t)$ as follows:

$$\begin{aligned}\theta(t) &= \max_{x \in R}[\phi(t) + (x - t)' \nabla\phi(t)] \\ &= \max_{x \in R}[H(x, t)].\end{aligned} \tag{5}$$

Now $\theta(t)$ is an unconstrained function of t and if we can relate the behavior of $\theta(t)$ to the initial optimization problem we may find this of value in solving the initial optimization problem.

This paper establishes a minimax theorem, which may be used to facilitate the solution of the initial optimization problem, and will indicate how this optimization procedure might be tackled.

The basic theorem is that

$$\min_{t\in E_n} \max_{x\in R}[H(x, t)] = \max_{x\in R} \min_{t\in E_n}[H(x, t)], \tag{6}$$

or, equivalently,

$$\min_{t\in E_n}[\theta(t)] = \max_{x\in R}[\phi(x)]. \tag{7}$$

In the case when the optimum unconstrained solution of $\phi(x)$ is already within R, the minimax theorem is trivial.

From Lemma 1, we see that

$$\max_{x\in R} \min_{t}[H(xt)] \leqslant \min_{t} \max_{x\in R}[H(x, t)]. \tag{8}$$

Now let t^u be the unconstrained optimum of $\theta(t)$.

Then $\nabla\phi(x^*) = 0$ and we also have

$$\begin{aligned} \min_{t} \max_{x\in R}[H(x, t)] &\leqslant \max_{x\in R}[H(x, t^u)] \\ &= \theta(t^u) \\ &\leqslant \theta(x^*) \\ &= \phi(x^*) = \max_{x\in R} \min_{t}[H(x, t)], \end{aligned} \tag{9}$$

if x^* maximizes $\phi(x)$, $x \in R$.

In this case, the computational implications of the minimax theorem are zero, since we simply find the unconstrained optimum and check that it lies in R. If it does, we need go no further. If it does not, then the computational procedures implicit in what follows may be used.

The theoretical development is split into two parts. First, we consider the case when the boundary of R is strictly convex. Then we consider the case when the boundary of R is simply convex, thus allowing us to include optimization subject to linear constraints.

In what follows we assume R bounded, thus ensuring finite x's, t's, and also $\phi(x)$ is bounded, thus ensuring $\theta(t)$ is bounded. This in no way loses the generality of application, since we can make these bounds as large as we wish.

2. Theoretical Development

In what follows, we define $x(t)$ to be a specified value of x which minimizes $\theta(t)$ for a given t.

In the case when R is strictly convex, and $\nabla\phi(t) = 0$, $x(t)$ will be a unique point. When R is convex, $x(t)$ is defined equal to $x_0(t)$ in Lemma 2[I].

2.1. *Case when the Boundary of R is strictly Convex*

We first of all prove a series of 8 lemmas and end with the basic theorem. By strict convexity we mean that if x_1 and $x_2 \in R$, then $\lambda x_1 + (1 - \lambda)\, x_2$ is an interior point of R for $0 < \lambda < 1$.

Lemma 1.

$$\max_{x \in R} \min_t [H(x, t)] \leqslant \min_t \max_{x \in R} [H(x, t)]. \tag{10}$$

Proof.

$$\begin{aligned} \min_t [H(x^0, t)] &\leqslant H(x^0, t^0) \\ &\leqslant \max_{x \in R} [H(x, t^0)] \qquad \text{for all } t^0 \text{ and all } x^0 \in R. \end{aligned} \tag{11}$$

Then,

$$\begin{aligned} \max_{x^0 \in R} \min_t [H(x^0, t)] &\leqslant \max_{x \in R} [H(x, t^0)], \qquad \text{for all } t^0 \\ &\leqslant \min_{t^0} \max_{x \in R} [H(x, t^0)]. \end{aligned} \tag{12}$$

This is essentially the required result, dropping the suffixes.

Lemma 2. *Let $a'(t)\, x = b(t)$ be a hyperplane tangent to R at $x(a(t))$. Let $a(t)$ be differentiable, and let the boundary of R be strictly convex. Assume $a(t) \neq 0$.*

Then,

$$\varDelta x'(a(t)) \cdot a(t) \sim 0 \,\|\, \varDelta t \,\|^2, \tag{13}$$

where

$$\| z \| = \max_i | z_i | .$$

Proof. Let $d(x)$ be the distance from x to $a'(t)$. $x = b(t)$. Then,

$$d(x + \varDelta x) = d(x) + \varDelta x' \,\nabla d(x) + 0 \,\|\, \varDelta x \,\|^2. \tag{14}$$

Now at $x = x(a(t))$, we have $d(x) = 0$ and it is a minimum value in the region of $x = x(a(t))$. Thus $\nabla d(x) = 0$ at this point.

Also,

$$\| \Delta x(a(t))\|^2 \sim 0 \| \Delta t \|^2. \tag{15}$$

Finally,

$$\Delta x'(a(t)) \cdot a(t) = d(x(a(t)) + \Delta x(a(t))). \tag{16}$$

Thus the lemma is proved.

LEMMA 3.

$$\Delta \theta(t) = (x(t) - t)' \nabla^2 \phi(t) \Delta t + 0 \| \Delta t \|^2. \tag{17}$$

For the time being we shall assume that we encounter no t for which $\nabla\phi(t) = 0$ and all the Lemmas 3–8 and the theorem make this assumption. The reason for this is that when $\nabla\phi(t) = 0$, there is no unique optimizing value of x for $H(x, t)$. Since we shall be encountering the same problem when R is not strictly convex, we assume, for the time being, that $\nabla\phi(t) \neq 0$. In Section 2.3 we return to this point.

Proof.

$$\begin{aligned} \Delta \theta(t) &= \Delta \phi(t) + \Delta (x(t) - t)' \nabla \phi(t) \\ &= \Delta t' \nabla \phi(t) + \Delta x'(t) \nabla \phi(t) - \Delta t' \nabla \phi(t) + (x(t) - t)' \nabla^2 \phi(t) \Delta t \\ &\quad + 0 \| \Delta t \|^2. \end{aligned} \tag{18}$$

From Lemma 2, we have, with $a(t) \equiv \nabla' \phi(t)$,

$$\Delta x'(t) \nabla \phi(t) = 0 \| \Delta t \|^2. \tag{19}$$

Thus,

$$\Delta \theta(t) = (x(t) - t)' \nabla^2 \phi(t) \Delta t + 0 \| \Delta t \|^2. \tag{20}$$

LEMMA 4. *If t^* minimizes $\theta(t)$, then $x(t^*) = t^*$.*

Proof. In Lemma 3, let $\Delta t = \lambda(x - t^*)$. λ small and positive. Then,

$$\Delta \theta(t^*) = \lambda(x(t^*) - t^*)' \nabla^2 \phi(t^*) (x(t^*) - t^*) + 0(\lambda^2). \tag{21}$$

Since $\phi(t)$ is strictly concave, $\nabla^2\phi(t^*)$ is negative definite. Thus we can decrease $\theta(t)$ below the value of $\theta(t^*)$ unless $x(t^*) = t^*$. Since $\theta(t^*)$ is the lowest value of $\theta(t)$, we must have $x(t^*) = t^*$.

LEMMA 5. *If x^* maximizes $\phi(x)$, subject to $x \in R$, then x^* is one value for $x(x^*)$.*

Proof. Suppose $x(x^*) = x^{**}$, and x^* is not a value of $x(x^*)$.

Then, if $0 \leqslant \lambda \leqslant 1$, we have

$$\begin{aligned}
&\phi(\lambda x^{**} + (1-\lambda)\, x^*) - \phi(x^*) \\
&\quad = \phi(x^* + \lambda(x^{**} - x^*)) - \phi(x^*) \\
&\quad = \lambda(x^{**} - x^*)' \, \nabla\phi(x^*) + 0(\lambda^2) \qquad \text{if } \lambda \text{ is small enough.}
\end{aligned} \tag{22}$$

Now

$$0 < H(x^{**}, x^*) - H(x^*, x^*) = (x^{**} - x^*)' \, \nabla\phi(x^*). \tag{23}$$

Thus, $\phi(\lambda x^{**} + (1-\lambda)\, x^*) > \phi(x^*)$ for some sufficiently small λ. This contradicts the fact that x^* maximizes $\phi(x)$, for $x \in R$, since $\lambda x^{**} + (1-\lambda)\, x^* \in R$ and since R is convex.

LEMMA 6.

$$\theta(t^*) = \phi(x^*) = \phi(x(t^*)). \tag{24}$$

Proof.

$$\begin{aligned}
\phi(x^*) &\leqslant \theta(t^*) && \text{(Lemma 1)} \\
&\leqslant \theta(x^*) \\
&= H(x^*, x^*) && \text{(Lemma 5)} \\
&= \phi(x^*).
\end{aligned} \tag{25}$$

Also,

$$\begin{aligned}
\phi(x(t^*)) &= \phi(t^*) && \text{(Lemma 4)} \\
&= \theta(t^*) && \text{(Lemma 4)} \\
&= \phi(x^*).
\end{aligned} \tag{26}$$

LEMMA 7. *If t^0 is any local minimum of $\theta(t)$, then*

$$\phi(t^0) = \theta(t^0) = \phi(x^*) \qquad \textit{and} \qquad x(t^0) = t^0.$$

Proof. Following the same procedure as in Lemma 4, if t^0 is a local minimum of $\theta(t)$, we must still have $x(t^0) = t^0$.

Then,

$$\theta(t^0) = \phi(t^0) \leqslant \phi(x^*). \tag{27}$$

Also,

$$\theta(t^0) = H(t^0, t^0) \geqslant \phi(x^*) \qquad \text{(Lemma 1)}. \tag{28}$$

Hence,

$$\theta(t^0) = \phi(x^*). \tag{29}$$

LEMMA 8. *$\theta(t)$ has only one local minimum.*

Proof. Since $\phi(x)$ is strictly convex, it has only one maximum point x^*. From Lemma 7, if t^0 is a local minimum of $\theta(t)$, we must have therefore, $t^0 = x^*$.

THEOREM.

$$\max_{x \in R} \min_t [H(x, t)] = \min_t \max_{x \in R} [H(x, t)], \tag{30}$$

and if t^* is the, unique, local minimum of $\theta(t)$, then $t^* = x^*$.

Proof. This follows directly from Lemmas 7 and 8.

2.2. *Case when the Boundary of R is simply Convex*

We prove two lemmas which are the counterparts of Lemmas 2, 3, 4, with Lemmas 3 and 4 being combined because of the special analysis needed. The remaining results are identical.

By simple convexity we mean that if x_1 and $x_2 \in R$, then

$$\lambda x_1 + (1 - \lambda) x_2 \in R \qquad \text{for } 0 \leqslant \lambda \leqslant 1.$$

LEMMA 2^{I}. *Let $a'(t)\, x = b(t)$ be a tangent hyperplane to R, with the common region S containing more than one point of R. Let $a(t)$ be differentiable, $\nabla a(t)$ be strictly negative definite, where $\nabla a(t) = (\delta a(t)/\delta t_i)$. Then there exists an $x_0 \in S$ such that*

$$(x - x_0)' \nabla a(t) (x_0 - t) \leqslant 0, \qquad \textit{for all } x \in S. \tag{31}$$

Proof. The lemma is clearly true if $t \in S$, and we can select $x_0 = t$. Suppose $t \notin S$. Since $\nabla a(t)$ is negative definite, there exists a nonsingular transformation matrix T such that $T' \nabla a(t)\, T \equiv -I$ where I is the unit $n \times n$ matrix.

Let $\xi = Tx^{-1}$, $\tau = T^{-1}t$, and let $S^{-1} \equiv T^{-1}S$.

Since $t \notin S$, then $\tau \notin S^{-1}$.

Then we have

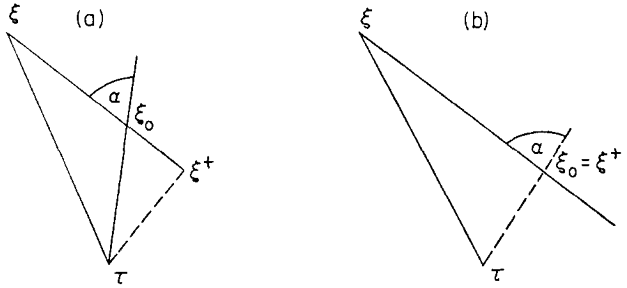
$$(x - x_0)' \nabla a(t) (x_0 - t) = -(\xi - \xi_0)' (\xi_0 - \tau). \tag{32}$$

Now choose ξ_0 to minimize $\xi \in S^{-1}[(\xi - \tau)' (\xi - \tau)]$.

Let the perpendicular from τ to the line joining ξ and ξ_0 meet this line at the point ξ^+.

Then either ξ and ξ_0 are on one side of the perpendicular (Fig. 1a) or $\xi_0 = \xi^+$ (Fig. 1b). This follows because S, and hence S^{-1}, is a convex set, thus preventing ξ_0 and ξ from being on different sides of the perpendicular, since, if they were, every point joining ξ and ξ_0 would also be in S^{-1}, and one of these points would be ξ^+ and provide a smaller value of $(\xi - \tau)'(\xi - \tau)$.

In either case, $\cos(\alpha) \geqslant 0$ in Figs. 1a and b.



FIGURES 1a and b

If $\xi_0 \neq \tau$, $\cos(\alpha) = k(\xi - \tau)'(\xi_0 - \tau)$, where $k > 0$, and hence $(\xi - \tau)'(\xi_0 - \tau) \geqslant 0$. If $\xi_0 = \tau$, we trivially get

$$(\xi - \tau)'(\xi_0 - \tau) = 0.$$

Thus the lemma is proved.

LEMMA 2^{II}. *If S contains only one point, x_0, then lemma 2^{I} is trivially true.*

LEMMA 4^{I}. *If t^* minimizes $\theta(t)$, then $x(t^*) = t^*$.*

Proof. From the general theory of Linear Programming, it is known that if $a(t)'(x - t)$ takes its maximum value at any point in S, then $\exists\epsilon > 0$ such that $a'(t + \Delta t)(x - t - \Delta t)$ takes its maximum value at at least one point in S providing $\| \Delta t \| \leqslant \epsilon$.

Now let $a(t) \equiv \nabla\phi(t)$ and let $x(t + \Delta t)$ be the optimum $x \in S$ for $\theta(t + \Delta t)$. Let $x_0(t)$ be chosen as per Lemmas 2^{I} or 2^{II}.

Let t^* minimize $\theta(t)$.

Then,

$$\theta(t^* + \Delta t) = \phi(t^* + \Delta t) + (x(t^* + \Delta t) - t^* - \Delta t)'\nabla\phi(t^* + \Delta t), \tag{33}$$

$$\theta(t^*) = \phi(t^*) + (x_0(t^*) - t^*)'\nabla'\phi(t^*). \tag{34}$$

$$\begin{aligned}\Delta\theta(t^*) = {} & (x(t^* + \Delta t) - x_0(t^*))'\nabla\phi(t^*) + (x(t^* + \Delta t) - t^*)' \\ & \times \nabla^2\phi(t^*)\Delta t + 0\| \Delta t \|^2.\end{aligned} \tag{35}$$

Now we have

$$\begin{gathered} x(t^* + \Delta t)'\nabla\phi(t^*) = x_0(t^*)'\nabla\phi(t^*), \\ (\text{since } x(t^* + \Delta t) \text{ and } x_0(t^*) \in S). \end{gathered} \tag{36}$$

Also,

$$
\begin{aligned}
(x(t^* + \Delta t) - t^*)' \, \nabla^2\phi(t^*) \, \Delta t & \\
&= \lambda(x_0(t^*) - t^*)' \, \nabla^2\phi(t^*) \, (x_0(t^*) - t^*) \\
&\quad + \lambda(x(t^* + \Delta t) - x_0(t^*))' \, \nabla^2\phi(t^*) \, (x_0(t^*) - t^*) \\
&\leqslant \lambda(x_0(t^*) - t^*)' \, \nabla^2\phi(t^*) \, (x_0(t^*) - t^*) \text{ using Lemmas } 2^{\mathrm{I}} \text{ or } 2^{\mathrm{II}} \\
&< 0
\end{aligned}
\tag{37}
$$

if

$$x_0(t^*) \neq t^* \qquad \text{and} \qquad 0 < \lambda \, \| \, x_0(t^*) - t^* \, \| \leqslant \epsilon.$$

Thus $\Delta\theta(t^*) < 0$ if $x_0(t^*) \neq t^*$, and the lemma is proved.

The remaining Lemmas 5–8 now follow automatically and the main theorem also holds true.

2.3. *Case when* $\nabla\phi(t) = 0$

In the case when $\nabla\phi(t) = 0$, then all $x \in R$ maximize $H(x, t)$. Let us now define S to be identical with R, and choose $x_0(t)$ to maximize $(x - t)' \, \nabla^2\phi(t) \, (x - t)$. The results of Section 2.2 then hold, and the Lemmas 5–8 then follow automatically.

3. Practical Implications

Suppose we have reached a stage when we have a particular value of t and $S(t)$ is the set of all $x \in R$ which maximize $H(x, t)$.

We let $x_0(t) \in S$ be a value of $x \in S$ which maximizes

$$(x - t)' \, \nabla^2\phi(t) \, (x - t).$$

If $x_0(t) = t$, then $x_0(t)$ is the unique optimum solution to our initial optimization problem.

If $x_0(t) \neq t$, we set $\Delta t = \lambda(x_0(t) - t)$ and find $x(t + \Delta t)$ to minimize $\theta(t + \Delta t)$. We may need to ensure that λ is small enough for $\theta(t + \Delta t)$ to be less than $\theta(t)$.

Clearly one might develop in due course different rules for determining λ.

Let us define

$$\delta(t, \lambda) = \theta(t + \lambda(x_0(t) - t)). \tag{38}$$

We know from inequality (37) that there is a nonzero $\bar{\lambda}$ such that for all λ less than or equal to this, $\delta(t, \lambda) < \delta(t, 0)$, $\lambda \neq 0$, unless we have already arrived at the optimum solution at t.

We can now use any procedure for locating the minimum, over λ, of the function $\delta(t, \lambda)$.

The convergence of the procedure follows from the facts that $\theta(t)$ has a unique local minimizing point (Lemma 8), identical with the unique maximizing point, in R, of $\phi(x)$; we can always reduce $\theta(t)$ unless we have reached this point (inequality 36); and an assumption that the minimum and maximum values of $\theta(t)$ and $\phi(x)$ ($x \in R$) are bounded.

Where R is strictly convex, S contains only one point and the next stage is easy.

When R is simply convex, S may contain more than one point, we face a quadratic programming problem at some stages; but, even in this case, most of the sequences of sets S arising in the calculations will consist of one point.

If R is the region determined by linear constraints, most of our optimum $x(t)$, in a given series of calculations, will be single vertices of the convex polyhedron R, and we will have $x_0(t + \Delta t) \equiv x_0(t)$ until Δt is sufficiently large. When Δt becomes sufficiently large $S(t + \Delta t)$ will eventually contain a further point to $x_0(t)$, and hence all points on the segment joining $x_0(t)$ to this point.

In such cases, the relative costs for the nonbasic variables of the implicit linear program will be a known linear function of

$$\phi(t + \Delta t) = \phi(t + \lambda(x_0 - t));$$

and we simply increase λ until one of these becomes zero. Usually only one will become zero, although cases will arise when more than one become zero.

In the former case, if x_1 is the next vertex to qualify for inclusion in the solution, then any point of $S(t + \Delta t)$ is a positive linear combination of x_0 and x_1. Then we have

$$\begin{aligned}(x - t - \Delta t)' \, \nabla^2\phi(t + \Delta t)\,(x - t - \Delta t) \\ = (\mu(x_0 - t - \Delta t) + (1 - \mu)\,(x_1 - t - \Delta t))' \, \nabla^2\phi(t + \Delta t) \qquad (39) \\ \times (\mu(x_0 - t - \Delta t) + (1 - \mu)\,(x_1 - t - \Delta t)).\end{aligned}$$

Expression (39) is quadratic in μ and can easily be maximized over $0 \leqslant \mu \leqslant 1$ to give $x_0(t + \Delta t)$.

In the latter case, when several relative costs become zero, we have a more general quadratic program to solve, but this may contain few variables, since $S(t + \Delta t)$ may have only a few dimensions.

Alternate Proof[1]

A short, simple and direct proof of the author's minimax theorem is the following:

In Ref. [4, p. 96, Theorem (13B)][1], it is proved that Eq. (6) of the paper holds if and only if $H(x, t)$ has a saddle point. It is quite easy to establish the existence of a saddle point for $H(x, t)$.

Let x_0 be a point of R such that $\phi(x_0) \geqslant \phi(x)$ for all $x \in R$. Let $t_0 = x_0$. In order to show that (x_0, t_0) is a saddle point, it suffices to show that

$$\begin{aligned} &\phi(t_0) + (x - t_0)' \nabla\phi(t_0) \\ &\quad \leqslant \phi(t) + (x_0 - t)' \nabla\phi(t) \qquad \text{for all } x \in R \text{ and all } t \in E^n. \end{aligned} \tag{A-1}$$

Since $t_0 = x_0$, the inequality (A-1) can be rewritten as

$$\phi(t_0) + (x - x_0)' \nabla\phi(x_0) \leqslant \phi(t) + (t_0 - t)' \nabla\phi(t). \tag{A-2}$$

Note that (A-2) follows from the following two inequalities:

$$\phi(t_0) - \phi(t) \leqslant (t_0 - t)' \nabla\phi(t) \qquad \text{for all } t \in E^n. \tag{A-3}$$

$$(x - x_0) \nabla\phi(x_0) \leqslant 0 \qquad \text{for all } x \in R. \tag{A-4}$$

The inequality (A-3) holds since ϕ is concave (it is not necessary that ϕ be strictly concave). The inequality (A-4) holds since ϕ is concave, R is convex, and since $\phi(x_0) \geqslant \phi(x)$ for all $x \in R$; that is, note that $(x - x_0)' \cdot \nabla\phi(x_0)$ is the directional derivative of ϕ at x_0 in the direction of $(x - x_0)$. By the definition of the directional derivative, we have

$$(x - x_0)' \cdot \nabla\phi(x_0) = \lim_{\alpha \to 0} = \frac{\phi(x_0 + \alpha(x - x_0)) - \phi(x_0)}{\alpha}. \tag{A-5}$$

Since R is convex, we have $x_0 + \alpha(x - x_0) \in R$ for all $\alpha \in (0, 1)$. Since $\phi(x_0 + \alpha(x - x_0)) - \phi(x_0) \leqslant 0$ for all $\alpha \in (0, 1)$, we get from (A-5) the inequality (A-4). This completes the proof of the minimax theorem.

Acknowledgments

Apart from the intrinsic mathematical aspects of a general nature used in this development, this development is not a follow-up on existing developments in the area of mathematical programming. It is related to the "projected gradient" and "reduced gradient" methods referred to be Wolfe [1] in that it uses gradient ideas and the linear programming simplex method and incidental relative costs in getting the solution. The real inspiration came from the Kuhn–Tucker minimax theorem [2],

[1] Dr. H. Staford, Division of Applied Mechanics, University of California, Berkeley.

and, in particular, in the presentation of this and the corresponding duality theory from Lasdon [3].

I also wish to acknowledge useful points raised by the referees and am indebted to one of them[1] who produced a very simple proof of the minimax theorem which is appended.

References

1. P. Wolfe, Methods of nonlinear programming, *in* "Recent Advances in Mathematical Programming" (R. Graves and P. Wolfe, Eds.), McGraw-Hill, New York, 1963.
2. H. Kuhn and A. Tucker, "Non-Linear Programming, Proceedings of Second Berkeley Symposium on Mathematical Statistics and Probability," Univ. of California Press, Berkeley, CA, 1951.
3. L. Lasdon, Duality and decomposition in mathematical programming, *IEEE Trans. Systems Sci. Cybernetics* **SSC4** (2) (1968).
4. John Von Neumann and O. Morgenstern, "Theory of Games and Economic Behavior," Princeton Univ. Press, Princeton, NJ, 1944; Science edition, Wiley, New York, 1964.